



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/737,190

Source: 07PE

Date Processed by STIC: 7/25/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
- 2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).

Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:

<http://www.uspto.gov/web/offices/pac/checker>

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER: 09/737,190

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 Wrapped Nucleic
 Wrapped Aminos

 The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor **after** creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2 Invalid Line Length

 The rules require that a line **not exceed** 72 characters in length. This includes white spaces.

3 Misaligned Amino
 Numbering

 The numbering under each 5th amino acid is misaligned. **Do not** use tab codes between numbers; use **space characters**, instead.

4 Non-ASCII

 The submitted file was **not** saved in ASCII(DOS) text, as required by the Sequence Rules. **Please ensure your subsequent submission is saved in ASCII text.**

5 Variable Length

 Sequence(s) _____ contain n's or Xaa's representing more than one residue. **Per Sequence Rules, each n or Xaa can only represent a single residue.** Please present the **maximum** number of each residue having variable length and indicate in the <220>.<223> section that some may be missing.

6 PatentIn 2.0
 "bug"

 A "bug" in PatentIn version 2.0 has caused the <220>.<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>.<223> section to the subsequent amino acid sequence. **This applies to the mandatory <220>.<223> sections for Artificial or Unknown sequences.**

7 Skipped Sequences
 (OLD RULES)

 Sequence(s) _____ missing. If intentional, please insert the following lines for **each** skipped sequence:
 (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
 (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped

 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to **include** the skipped sequences.

8 Skipped Sequences
 (NEW RULES)

 Sequence(s) _____ missing. If **intentional**, please insert the following lines for **each** skipped sequence.
 <210> sequence id number
 <400> sequence id number
 000

9 Use of n's or Xaa's
 (NEW RULES)

 Use of n's and/or Xaa's have been detected in the Sequence Listing.
 Per 1.823 of Sequence Rules, use of <220>.<223> is **MANDATORY** if n's or Xaa's are present.
 In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10 Invalid <213>
 Response

 Per 1.823 of Sequence Rules, the **only valid** <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>.<223> section is **required** when <213> response is Unknown or is Artificial Sequence

11 Use of <220>

 Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses.
 Use of <220> to <223> is **MANDATORY** if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
 (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

12 PatentIn 2.0
 "bug"

 Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13 Misuse of n

 n can only be used to represent a single nucleotide in a nucleic acid sequence. N is **not** used to represent any value not specifically a nucleotide.

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/737,190

DATE: 07/25/2001

TIME: 13:05:30

Input Set : A:\14043.asc

Output Set: N:\CRF3\07252001\I737190.raw

ERRORED SEQUENCES

351 <210> SEQ ID NO: 2
 352 <211> LENGTH: 1334
 353 <212> TYPE: DNA
 354 <213> ORGANISM: Streptococcus anginosus
 356 <400> SEQUENCE: 2
 357 gaacgggtga gtaacgcgta ggtaaacctgc ctattagagg gggataacta ttggaaacga 60
 359 tagctaatac cgcataaacag tatgtAACAC atgttagatg cttgaaaAGAT gcaATTGcat 120
 361 cgctAGTAGA tggacCTGCG ttgttattAGC tagtaggtAG ggtaaAGGCC tacCTAGGC 180
 363 acgatacata gccgacCTGA gagGGTgATC ggCCACACTG ggACTgAGAC acggCCCAGA 240
 365 ctcctacGGG aggCAGCAgT agggAAATCTT CGGCAATGGG gggAACCTGT accgAGCAAC 300
 367 gccgCGTgAG tgaAGAAAGT ttTCGGATCG taaAGCTGT ttgttaAGGA aGAACGAGTG 360
 E--> 369 tgagaatggA aagtTCATAc TGTGACGGTA CTTAACCAgA aAGGGACGGC t~~n~~actacGTG 420 → see item 9
 E--> 371 ccagcAGCCG CGGTAATAcG taggtCCoNA gcgttGTCG gatttATTGG gCGTAAAGCG 480
 373 agcgcAGGCG gttAGAAAAG tctGAAGTGA aaggcAGTGG ctaaccATT gtaggCTTT 540
 375 gaaACTGTtT aacttgAGtG cagaAGGGGA gAGTGGAAATT ccatgtGTAG CGGTGAAATG 600
 377 cgtAGATA TGGAGGAACA CCggTGGCGA aAGCggCTCT CTGTCtGTA actgACGCTG 660
 E--> 379 aggCTCgAAa GCGTGGGGAG CGAACAGGAT tagATACCT ~~t~~AGTGTCCAC gCGTAAACG 720
 381 atgagtGCTA ggtGTTGGT CCTTCCGGG actcAGTGCc gcaGCTAACG cattaAGCAC 780
 383 tccgcCTGGG gagtACGACC gcaAGGTTGA aactCAAAGG aattGACGGG ggCCGcacAA 840
 E--> 385 gCGGTGGAGC atgtnGTTTA attCGAAGmA acgCGAAGAA CCTTACCAgg tCTTGACATC 900
 E--> 387 ccgatGCTnT ttctAGAGAT aggaAGTTc ttCGGAACAT CGGTGACAGG tGGTGCATGG 960
 E--> 389 ttGTCGTcAG CTCGTGTcGT gagATGTTGG ttAAAGTCC GCAACGAGCG caACCCTnT 1020
 E--> 391 tGTTAGTTGc catCATTAAg ttGGGCACTC tagCGAGACT gCCGTAATnT aaccCGGAGGA 1080
 E--> 393 aggtGGGGAT gacGTCAAAt catCATGCC CTTATGACCT ~~t~~GGTACACAA CGTGTACAA 1140
 395 tggCTGTGAC aacgAGTCGc aAGCCGGTGA CGGCAAGCTA ATCTCTGAAA GCCAGTCTCA 1200
 E--> 397 gttcGGGATTG tagGTGCAA CTCGCTnTCA tGAAGTGGTA ATCGCTAGTA ATCGCGGATC 1260
 E--> 399 agcACGCCGc ggtGAATAcG ttCCCGGGCC ttGTACACAC CGC~~n~~GTCAC ACCACGAGAG 1320
 401 ttGTAACAC CGGA 1334
in Error
Summary Sheet
(erroneous throughout
Seq. 2)

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/737,190

DATE: 07/25/2001

TIME: 13:05:31

Input Set : A:\14043.asc

Output Set: N:\CRF3\07252001\I737190.raw

L:369 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:2

M:340 Repeated in SeqNo=2